



University of  
Zurich<sup>UZH</sup>

Zurich Open Repository and  
Archive

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2020

---

## How linear reinforcement affects Donsker's theorem for empirical processes

Bertoin, Jean

Abstract: A reinforcement algorithm introduced by Simon (Biometrika 42(3/4):425–440, 1955) produces a sequence of uniform random variables with long range memory as follows. At each step, with a fixed probability  $p \in (0,1)$ ,  $U^{n+1}$  is sampled uniformly from  $U^1, \dots, U^n$ , and with complementary probability  $1-p$ ,  $U^{n+1}$  is a new independent uniform random variable. If  $p > 1/2$ , and if a further rescaling is needed when  $p > 1/2$  and the limit is then a bridge with exchangeable increments and discontinuous paths, called elephant random walk, and more generally step reinforced random walks.

DOI: <https://doi.org/10.1007/s00440-020-01001-9>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-190455>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Bertoin, Jean (2020). How linear reinforcement affects Donsker's theorem for empirical processes. Probability Theory and Related Fields, 178:1173-1192.

DOI: <https://doi.org/10.1007/s00440-020-01001-9>



# How linear reinforcement affects Donsker's theorem for empirical processes

Jean Bertoin<sup>1</sup>

Received: 18 May 2020 / Revised: 4 September 2020 / Accepted: 9 September 2020  
© The Author(s) 2020

## Abstract

A reinforcement algorithm introduced by Simon (Biometrika 42(3/4):425–440, 1955) produces a sequence of uniform random variables with long range memory as follows. At each step, with a fixed probability  $p \in (0, 1)$ ,  $\hat{U}_{n+1}$  is sampled uniformly from  $\hat{U}_1, \dots, \hat{U}_n$ , and with complementary probability  $1 - p$ ,  $\hat{U}_{n+1}$  is a new independent uniform variable. The Glivenko–Cantelli theorem remains valid for the reinforced empirical measure, but not the Donsker theorem. Specifically, we show that the sequence of empirical processes converges in law to a Brownian bridge only up to a constant factor when  $p < 1/2$ , and that a further rescaling is needed when  $p > 1/2$  and the limit is then a bridge with exchangeable increments and discontinuous paths. This is related to earlier limit theorems for correlated Bernoulli processes, the so-called elephant random walk, and more generally step reinforced random walks.

**Keywords** Donsker's theorem · Empirical process · Linear reinforcement · Bridges with exchangeable increments

**Mathematics Subject Classification** 60F17 · 62G30

## 1 Introduction

A classical result of Glivenko and Cantelli in 1933 states that the sequence of empirical distribution functions associated to i.i.d. copies of some real random variable converges uniformly to its cumulative distribution function, almost surely. Nearly 20 years later, Donsker determined the asymptotic behavior of the fluctuations; let us recall his result. Let  $U_1, U_2, \dots$  be i.i.d. uniform random variables on  $[0, 1]$ ; then the sequence of

---

Dedicated to the memory of Harry Kesten, for the deep mathematics he gave us.

---

✉ Jean Bertoin  
[jean.bertoin@math.uzh.ch](mailto:jean.bertoin@math.uzh.ch)

<sup>1</sup> Institute of Mathematics, University of Zurich, Zurich, Switzerland

(uniform) empirical processes,

$$G_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{U_i \leq x} - x), \quad x \in [0, 1],$$

converges in distribution as  $n \rightarrow \infty$  in the sense of Skorokhod towards a Brownian bridge  $(G(x))_{0 \leq x \leq 1}$ . The purpose of the present work is to analyze how Donsker's Theorem is affected by an elementary random reinforcement algorithm that we shall now describe.

Consider a sequence  $\varepsilon_2, \varepsilon_3, \dots$  of i.i.d. Bernoulli variables with fixed parameter  $p \in (0, 1)$ . These variables determine when repetitions occur, in the sense that the  $n$ -th step of the algorithm is a repetition if  $\varepsilon_n = 1$ , and an innovation if  $\varepsilon_n = 0$ . For every  $n \geq 2$ , let also  $v(n)$  be a uniform random variable on  $\{1, \dots, n-1\}$  such that  $v(2), v(3), \dots$  are independent; these variables specify which of the preceding items is copied when a repetition occurs. More precisely, we set  $\varepsilon_1 = 0$  for definitiveness and construct recursively a sequence of random variables  $\hat{U}_1, \hat{U}_2, \dots$  by setting

$$\hat{U}_n = \begin{cases} \hat{U}_{v(n)} & \text{if } \varepsilon_n = 1, \\ U_{i(n)} & \text{if } \varepsilon_n = 0, \end{cases}$$

where

$$i(n) = \sum_{j=1}^n (1 - \varepsilon_j) \quad \text{for } n \geq 1$$

denotes the total number of innovations after  $n$  steps. We always assume without further mention that the sequences  $(v(n))_{n \geq 2}$ ,  $(\varepsilon_n)_{n \geq 2}$ , and  $(U_j)_{j \geq 1}$  are independent.

This random algorithm has been introduced by Simon [28], who singled out in this setting a remarkable one-parameter family of power tail distributions on  $\mathbb{N}$  that arise in a variety of data. Nowadays, Simon's algorithm should be viewed as a linear reinforcement procedure, in the sense that, provided that  $i(n) \geq j$  (i.e. the variable  $U_j$  has already appeared at the  $n$ -th step of the algorithm), the probability that  $U_j$  is repeated at the  $(n+1)$ -th step is proportional to the number of its previous occurrences. In this direction, we refer henceforth to the parameter  $p$  of the Bernoulli variables  $\varepsilon_n$  as the reinforcement parameter.

Obviously, each variable  $\hat{U}_n$  has the uniform distribution on  $[0, 1]$ ; note however that the reinforced sequence  $(\hat{U}_n)_{n \geq 1}$  is clearly not stationary, and is not exchangeable or even partially exchangeable either. We shall furthermore point out in Remark 2.3 that mixing also does not hold, and the study of the empirical distribution functions for the reinforced sequence does not fit the framework developed for weakly dependent processes (see for instance Chapter 7 in Rio [26] and references therein). It is easy to show that nonetheless, the conclusion of the Glivenko–Cantelli theorem is still valid in our setting:

**Proposition 1.1** *With probability one, it holds that*

$$\lim_{n \rightarrow \infty} \sup_{0 \leq x \leq 1} \left| \frac{1}{n} \sum_{i=1}^n (\mathbf{1}_{\hat{U}_i \leq x} - x) \right| = 0.$$

We are chiefly interested in the empirical processes  $\hat{G}_n$  associated to the reinforced sequence

$$\hat{G}_n(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{\hat{U}_i \leq x} - x), \quad x \in [0, 1].$$

Our main result shows that their asymptotic behavior as  $n \rightarrow \infty$  exhibits a phase transition for the critical parameter  $p_c = 1/2$ . Roughly speaking, when the reinforcement parameter  $p$  is smaller than  $1/2$ , then the analog of Donsker's Theorem holds for  $\hat{G}_n$ , except that the limit is now only proportional to the Brownian bridge. At criticality, i.e. for  $p = 1/2$ , convergence in distribution to the Brownian bridge holds after an additional rescaling of  $\hat{G}_n$  by a factor  $1/\sqrt{\log n}$ . Finally for  $p > 1/2$ ,  $n^{-p+1/2}\hat{G}_n$  now converges in probability and its limit is described in terms of some bridge with exchangeable increments and discontinuous sample paths.

Here is a precise statement, where the needed background on bridges with exchangeable increments in the supercritical case  $p > 1/2$  is postponed to the next section. Recall that  $G = (G(x))_{0 \leq x \leq 1}$  denotes the standard Brownian bridge. We further write  $\mathbb{D}$  for the space of càdlàg paths  $\omega : [0, 1] \rightarrow \mathbb{R}$  endowed with the Skorokhod topology (see Chapter 3 in [8] or Chapter VI in [17]). The notation  $\Rightarrow$  is used to indicate convergence in distribution of a sequence of processes in  $\mathbb{D}$ .

**Theorem 1.2** *The following convergences hold as  $n \rightarrow \infty$ :*

(i) *If  $p < 1/2$ , then*

$$\hat{G}_n \Rightarrow \frac{1}{\sqrt{1-2p}} G.$$

(ii) *If  $p = 1/2$ , then*

$$\frac{1}{\sqrt{\log n}} \hat{G}_n \Rightarrow G.$$

(iii) *If  $p > 1/2$ , then*

$$\lim_{n \rightarrow \infty} n^{-p+1/2} \hat{G}_n = B^{(p)} \quad \text{in probability on } \mathbb{D},$$

where  $B^{(p)} = (B^{(p)}(x))_{0 \leq x \leq 1}$  is the bridge with exchangeable increments described in the forthcoming Definition 2.4.

Our approach to Theorem 1.2 relies, at least in part, on a natural interpretation of Simon's algorithm in terms of Bernoulli bond percolation on random recursive trees. Specifically, we view  $\{1, 2, \dots, n\}$  as a set of vertices and each pair  $(j, v(j))$  for  $j = 2, \dots, n$  as edges; the resulting graph  $\mathbb{T}_n$  is known as the random recursive tree of size  $n$ , see Section 1.3 and Chapter 7 in [13]. We next delete each edge  $(j, v(j))$  if and only if  $\varepsilon_j = 0$ , in other words we perform a Bernoulli bond percolation with parameter  $p$  on  $\mathbb{T}_n$ . The percolation clusters are then given by subsets of indices at which the same variable is repeated, namely  $\{i \leq n : \hat{U}_i = U_j\}$  for  $j = 1, \dots, i(n)$ .

The sum of the squares of the cluster sizes

$$\mathcal{S}^2(n) = \sum_{j \geq 1} N_j(n)^2, \quad \text{with } N_j(n) = \#\{i \leq n : \hat{U}_i = U_j\}, \quad (1)$$

lies at the heart of the analysis of the reinforced empirical process  $\hat{G}_n$ . We shall see that its asymptotic behavior is given by

$$\mathcal{S}^2(n) \sim \begin{cases} n/(1-2p) & \text{if } p < 1/2, \\ n \log n & \text{if } p = 1/2, \\ n^{2p} R & \text{if } p > 1/2, \end{cases} \quad (2)$$

where  $R$  is some non-degenerate random variable. A rough explanation for the phase transition<sup>1</sup> in (2) is that the main contribution to  $\mathcal{S}^2(n)$  is due to a large number of microscopic clusters in the sub-critical case  $p < 1/2$ , and rather to a few mesoscopic clusters of size  $\approx n^p$  in the super-critical case  $p > 1/2$ . Even though (2) is not quite sufficient to establish Theorem 1.2, it is nonetheless a major step for its proof. More precisely, we shall rely on general results due to Kallenberg [19] on the structure of processes with exchangeable increments and explicit criteria for the weak convergence of sequences of the latter, and (2) appears as a key element in this setting.

The rest of this work is organized as follows. Section 2 is devoted to several preliminaries. We shall first present some key results due to Kallenberg on bridges with exchangeable increments and their canonical representations. We shall then recall a limit theorem for the numbers of occurrences  $N_j(n)$  which have been obtained in the framework of Bernoulli percolation on random recursive trees as well as the fundamental result of H.A. Simon about the frequency of microscopic clusters. Last, we shall compute explicitly the average  $\mathbb{E}(\mathcal{S}^2(n))$  using a simple recurrence identity and establish Proposition 1.1 on our way. Theorem 1.2 is then proven in Sect. 3. Finally, in Sect. 4, we discuss some connections between Theorem 1.2 and closely related results in the literature on step-reinforced random walks, including correlated Bernoulli processes and the so-called elephant random walk.

<sup>1</sup> Somehow, the fact that percolation on random recursive trees exhibits a phase transition with critical parameter  $p_c = 1/2$  bears a flavor similar to Kesten's celebrated achievement [21] for bond percolation on the square lattice. However, this resemblance is purely coincidental and superficial.

## 2 Preliminaries

### 2.1 Bridges with exchangeable increments

This section is adapted from Kallenberg [19], who rather uses the terminology interchangeable instead of exchangeable, and whose results are given in a more general setting. We also refer to [20] for many interesting properties of the sample paths of such processes.

Let  $B = (B(x))_{0 \leq x \leq 1}$  be a real valued process with càdlàg sample paths, and which is continuous in probability. We say that  $B$  has exchangeable increments if for every  $n \geq 2$ , the sequence of its increments  $B(k/n) - B((k-1)/n)$  for  $k = 1, \dots, n$ , is exchangeable, i.e. its distribution is invariant by permutations. We further say that  $B$  is a bridge provided that  $B(0) = B(1) = 0$  a.s.

According to Theorem 2.1 in [19], any bridge with exchangeable increments can be expressed in the form

$$B(x) = \sigma G(x) + \sum_{j=1}^{\infty} \beta_j (\mathbf{1}_{U_j \leq x} - x), \quad x \in [0, 1], \quad (3)$$

where  $\sigma$  is a nonnegative random variable,  $G$  a Brownian bridge,  $\beta = (\beta_j)_{j \geq 1}$  a sequence of real random variables with  $\sum_{j=1}^{\infty} \beta_j^2 < \infty$  a.s., and  $U = (U_j)_{j \geq 1}$  a sequence of i.i.d. uniform random variables, such that  $\sigma$ ,  $G$ ,  $\beta$  and  $U$  are independent. More precisely, if we further assume that the sequence  $(|\beta_j|)_{j \geq 1}$  is nonincreasing, which induces no loss of generality, then the series in (3) converges a.s. uniformly on  $[0, 1]$ .

One calls  $\sigma, \beta$  the canonical representation of  $B$ . Roughly speaking, (3) shows that the continuous part of  $B$  is a mixture of Brownian bridges (parametrized by the standard deviation), with mixture weights given by the random variable  $\sigma$ , and  $\beta$  describes the sequence of the jumps of  $B$ , each of them taking place uniformly at random on  $[0, 1]$  and independently of the others. The laws of  $\sigma$  and of  $\beta$  then entirely determine that of  $B$ .

The next lemma plays a key role in the proof of Theorem 1.2; it states two criteria that are tailored for our purposes, for the convergence of a sequence of bridges with exchangeable increments in terms of the canonical representations. The first part is a special case of Theorem 2.3 in [19]. The second part can be seen as an immediate consequence of the first and the well-known facts that Skorokhod's topology is metrizable and that convergence of a sequence of functions in  $\mathbb{D}$  to a continuous limit is equivalent to convergence for the supremum distance (see e.g. Section VI.1 in [17]); it can also be checked by direct calculation.

**Lemma 2.1** *For each  $n \geq 1$ , let  $B_n$  denote a bridge with exchangeable increments and canonical representation  $\sigma_n = 0$  and  $\beta_n = (\beta_{n,j})_{j \geq 1}$ .*

(i) *Suppose that*

$$\lim_{n \rightarrow \infty} \sup_{j \geq 1} |\beta_{n,j}| = 0 \text{ in probability,}$$

and that

$$\sum_{j=1}^{\infty} \beta_{n,j}^2 \implies \sigma^2$$

for some random variable  $\sigma \geq 0$ . Then there is the convergence in distribution

$$B_n \implies \sigma G,$$

where  $G$  is a standard Brownian bridge independent of  $\sigma$ .

(ii) If

$$\lim_{n \rightarrow \infty} \sum_{j=1}^{\infty} \beta_{n,j}^2 = 0 \text{ in probability,}$$

then

$$\lim_{n \rightarrow \infty} \sup_{0 \leq x \leq 1} |B_n(x)| = 0 \text{ in probability.}$$

## 2.2 Asymptotic behavior of occurrences numbers

Recall that the reinforcement parameter  $p \in (0, 1)$  in Simon's algorithm is fixed; for the sake of simplicity, it will be omitted from several notations even though it always plays an important role.

Recall from (1) that for every  $j \in \mathbb{N}$ ,  $N_j(n)$  denotes the number of occurrences of the variable  $U_j$  up to the  $n$ -th step of the algorithm. Plainly  $N_j(n) = 0$  if and only if the number of innovations up to the  $n$ -th step is less than  $j$ , i.e.  $i(n) < j$ .

The starting point of our analysis is that the reinforced empirical process can be expressed in the form

$$\hat{G}_n(x) = \frac{1}{\sqrt{n}} \sum_{j=1}^{\infty} N_j(n) (\mathbf{1}_{U_j \leq x} - x), \quad x \in [0, 1]. \quad (4)$$

Hence  $\hat{G}_n$  is a bridge with exchangeable increments, with canonical representation 0 and  $\beta_n = (\beta_{n,j})_{j \geq 1}$ , where  $\beta_{n,j} = N_j(n)/\sqrt{n}$ . We aim to determine its asymptotic behavior as  $n \rightarrow \infty$  by applying Lemma 2.1. In this direction, the interpretation of Simon's algorithm as a Bernoulli bound percolation on a random recursive tree, as it has been sketched in the Introduction, enables us to lift from [1] the following result about the asymptotic behavior of mesoscopic clusters.

**Lemma 2.2** *The limit*

$$\lim_{n \rightarrow \infty} n^{-p} N_j(n) =: X_j^{(p)} \quad (5)$$

exists a.s. for every  $j \geq 1$ . For  $p > 1/2$ , there is furthermore the identity

$$\mathbb{E} \left( \sum_{j=1}^{\infty} (X_j^{(p)})^2 \right) = \frac{1}{(2p-1)\Gamma(2p)}.$$

**Proof** Simon's algorithm induces a natural partition  $\mathbb{N} = \bigsqcup_{j \geq 1} \Pi_j$  of the set of positive integers into blocks  $\Pi_j = \{k \in \mathbb{N} : \hat{U}_k = U_j\}$  which we can see as the result of a Bernoulli bond percolation with parameter  $p$  on the (infinite) random recursive tree. In this setting, we have  $N_j(n) = \#(\Pi_j \cap \{1, \dots, n\})$ , and the first claim of the statement has been observed in Section 3.2 of [1], right after the proof of Lemma 3.3 there<sup>2</sup>. Moreover Equation (3.4) there shows that for every  $q > 1/p$ , there is the identity

$$\mathbb{E} \left( \sum_{j=1}^{\infty} (X_j^{(p)})^q \right) = \frac{\Gamma(q)}{\Gamma(pq)} + \frac{q(1-p)\Gamma(q)}{(pq-1)\Gamma(pq)}.$$

Specializing this for  $q = 2$  yields our second claim.  $\square$

We write  $\mathbf{X}^{(p)} = (X_j^{(p)})_{j \geq 1}$ , where the  $X_j^{(p)}$  are defined by (5). It is known that  $X_1^{(p)}$  has the Mittag-Leffler distribution with parameter  $p$  (see Theorem 3.1 in [1] and also [24]); nonetheless the law of the whole sequence  $\mathbf{X}^{(p)}$  does not seem to have any simple expression (see Proposition 3.7 in [1]).

**Remark 2.3** Denote by  $J_* = \operatorname{argmax} \mathbf{X}^{(p)}$  the index at which the sequence  $(X_j^{(p)})_{j \geq 1}$  is maximal (it follows from Section 3.2 of [1] that  $J_* < \infty$  and is unique a.s.). We deduce from (5) that for every  $k \geq 1$ , the probability that  $U_{J_*}$  is the most repeated quantity in the sequence  $\hat{U}_k, \hat{U}_{k+1}, \dots, \hat{U}_{k+n}$  tends to 1 as  $n \rightarrow \infty$ . Hence the variable  $U_{J_*}$  is measurable with respect to the tail sigma-algebra

$$\hat{\mathcal{G}} = \cap_{k \geq 1} \sigma(\hat{U}_n : n \geq k).$$

On the other hand, it is readily checked that  $\mathbb{P}(J_* = 1) > 0$ ; as a consequence,  $\hat{\mathcal{G}}$  and  $U_1$  are not independent and the mixing property does not hold for the reinforced sequence  $(\hat{U}_n)_{n \geq 1}$ .

When  $p > 1/2$ , Lemma 2.2 enables us to view  $\mathbf{X}^{(p)}$  as a random variable with values in the space  $\ell^2(\mathbb{N})$  of square summable series, and this leads us to the following definition of the process  $B^{(p)}$  that appears as a limit in Theorem 1.2(iii).

**Definition 2.4** For  $p > 1/2$ , we define  $B^{(p)} = (B^{(p)}(x))_{0 \leq x \leq 1}$  as the bridge with exchangeable increments with canonical representation 0 and  $\mathbf{X}^{(p)}$ . That is

$$B^{(p)}(x) = \sum_{j=1}^{\infty} X_j^{(p)} (\mathbf{1}_{U_j \leq x} - x), \quad x \in [0, 1],$$

<sup>2</sup> The reinforcement parameter  $p$  here corresponds to  $e^{-t}$  in [1].



where  $\mathbf{U} = (U_j)_{j \geq 1}$  is a sequence of i.i.d. uniform variables, independent of  $\mathbf{X}^{(p)}$ .

We conclude this section recalling the key result of Simon [28] about the asymptotic frequency of microscopic percolation clusters. Note that the number of innovations up to the  $n$ -th step is approximately  $(1 - p)n$  for  $n \gg 1$ .

**Lemma 2.5** *For each  $k \geq 1$ , write*

$$C_k(n) := \#\{j \geq 1 : N_j(n) = k\},$$

*for the number of variables  $U_j$  which have occurred exactly  $k$  times at the  $n$ -th step of Simon's algorithm. Then*

$$\lim_{n \rightarrow \infty} \frac{C_k(n)}{(1 - p)n} = \frac{1}{p} B(k, 1 + 1/p) \quad \text{in probability for every } k \in \mathbb{N}, \quad (6)$$

where  $B$  denotes the beta function.

The right-hand side of (6) is a probability measure on  $\mathbb{N}$  which is known as the Yule-Simon distribution with parameter  $1/p$ . Actually, it is only proved in [28] that

$$\mathbb{E}(C_k(n)) \sim \frac{1 - p}{p} B(k, 1 + 1/p)n \quad \text{as } n \rightarrow \infty;$$

nonetheless the stronger statement (6) is known to hold; see e.g. Section 3.1 and more specifically Equation (3.10) in [25].

### 2.3 A first moment calculation

Recall that we want to apply Lemma 2.1 to investigate the asymptotic behavior of reinforced empirical processes. In this direction, (4) incites us to introduce for every  $n \geq 1$

$$\mathcal{S}^2(n) = \sum_{j=1}^{\infty} N_j(n)^2.$$

The proof of Theorem 1.2 will use the following explicit calculation for the expectation of this quantity, which already points at the same direction as (2).

**Lemma 2.6** *For every  $n \geq 1$ , we have*

$$\mathbb{E}(\mathcal{S}^2(n)) = \frac{\Gamma(n + 2p)}{\Gamma(n)} \sum_{i=1}^n \frac{\Gamma(i)}{\Gamma(i + 2p)}.$$

As a consequence, we have as  $n \rightarrow \infty$  that

$$\mathbb{E}(\mathcal{S}^2(n)) \sim \begin{cases} n/(1-2p) & \text{if } p < 1/2, \\ n \log n & \text{if } p = 1/2, \\ ((2p-1)\Gamma(2p))^{-1}n^{2p} & \text{if } p > 1/2. \end{cases}$$

**Proof** Write  $\mathcal{F}_n$  for the sigma-field generated by  $((\varepsilon_i, v(i)) : 2 \leq i \leq n)$ . Plainly,  $\sum_{j \geq 1} N_j(n) = n$ , and we see from the very definition of Simon's algorithm that

$$\begin{aligned} \mathbb{E}(\mathcal{S}^2(n+1) \mid \mathcal{F}_n) &= \mathcal{S}^2(n) + p \left( \frac{1}{n} \sum_{j \geq 1} N_j(n)(2N_j(n) + 1) \right) + (1-p) \\ &= (1 + 2p/n)\mathcal{S}^2(n) + 1. \end{aligned}$$

This yields the recurrence equation for the first moments

$$\mathbb{E}(\mathcal{S}^2(n+1)) = (1 + 2p/n)\mathbb{E}(\mathcal{S}^2(n)) + 1.$$

To solve the latter, we set  $a(n) = \Gamma(n+2p)/\Gamma(n)$ , so that  $a(n+1)/a(n) = 1 + 2p/n$ , and then

$$\mathbb{E}(\mathcal{S}^2(n+1))/a(n+1) = \mathbb{E}(\mathcal{S}^2(n))/a(n) + 1/a(n+1).$$

Since  $\mathcal{S}^2(1) = 1$ , we arrive at

$$\mathbb{E}(\mathcal{S}^2(n)) = a(n) \sum_{i=1}^n \frac{1}{a(i)},$$

which is the identity of our claim.

In turn, the estimate as  $n \rightarrow \infty$  in the statement follows immediately from the facts that  $\Gamma(n+2p)/\Gamma(n) \sim n^{2p}$ , and that when  $p > 1/2$ , one has

$$\begin{aligned} \sum_{i=1}^{\infty} \frac{\Gamma(i)}{\Gamma(i+2p)} &= \frac{1}{\Gamma(2p)} \sum_{i=1}^{\infty} B(i, 2p) \\ &= \frac{1}{\Gamma(2p)} \int_0^1 \left( \sum_{i=1}^{\infty} x^{i-1} \right) (1-x)^{2p-1} dx \\ &= \frac{1}{\Gamma(2p)} \int_0^1 (1-x)^{2p-2} dx \\ &= \frac{1}{(2p-1)\Gamma(2p)}. \end{aligned} \tag{7}$$

The proof is now complete.  $\square$

As a first application, we establish the reinforced version of the Glivenko–Cantelli theorem.

**Proof of Proposition 1.1** The proof is classically reduced to establishing the following reinforced version of the strong law of large numbers,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\hat{U}_i \leq x} = x \quad \text{a.s. for each } x \in [0, 1]. \quad (8)$$

Indeed, the almost sure convergence in (8) holds simultaneously for all dyadic rational numbers, and uniform convergence on  $[0, 1]$  then can be derived by a monotonicity argument *à la* Dini.

So fix  $x \in [0, 1]$  and set

$$\Sigma(n) = \sum_{i=1}^n \mathbf{1}_{\hat{U}_i \leq x} = \sum_{j=1}^{\infty} N_j(n) \mathbf{1}_{U_j \leq x}.$$

Clearly,  $\mathbb{E}(\Sigma(n)) = nx$ , and, by conditioning on  $\mathcal{F}_n$ , we get

$$\text{Var}(\Sigma(n)) = (x - x^2)\mathbb{E}(\mathcal{S}^2(n)).$$

From Lemma 2.6 and Chebychev’s inequality, we now see that we can choose  $r > 1$  sufficiently large such that

$$\sum_{k=1}^{\infty} \mathbb{P}(|\Sigma(k^r) - k^r x| > k^{r-1}) < \infty.$$

One concludes from the Borel–Cantelli lemma that (8) holds along the subsequence  $n = k^r$ , and the general case follows by another argument of monotonicity.  $\square$

### 3 Proof of Theorem 1.2

As its title indicates, the purpose of this section is to establish Theorem 1.2 in each of the three regimes.

#### 3.1 Subcritical regime $p < 1/2$

*Throughout this section, we assume that the reinforcement parameter satisfies  $p < 1/2$ . Our approach in the subcritical regime relies on the following strengthening of Lemma 2.5 (recall the notation there).*

**Lemma 3.1** *Define for every  $i \geq 1$*

$$c_i^{(p)} = \frac{(1-p)}{p} \text{B}(i, 1 + 1/p).$$

Then we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{i=1}^{\infty} i^2 \left| \frac{C_i(n)}{n} - c_i^{(p)} \right| \right) = 0.$$

**Proof** For each  $n = 1, 2, \dots$ , write  $\mathbf{C}(n) = (C_i(n))_{i \geq 1}$  and view  $\mathbf{C}(n)$  as a function on the space  $\Omega \times \mathbb{N}$  endowed with the product measure  $\mathbb{P} \otimes \#^2$ , where  $\#^2$  stands for the measure on  $\mathbb{N}$  which assigns mass  $i^2$  to every  $i \in \mathbb{N}$ . Consider an arbitrary subsequence excerpt from  $(\mathbf{C}(n))_{n \geq 1}$ . From Lemma 2.5 and an argument of diagonal extraction, we can construct a further subsequence, say indexed by  $\ell(n)$  for  $n = 1, 2, \dots$ , such that

$$\lim_{n \rightarrow \infty} \frac{\mathbf{C}(\ell(n))}{\ell(n)} = \mathbf{c}^{(p)} \quad (\mathbb{P} \otimes \#^2)\text{-almost everywhere,} \quad (9)$$

where  $\mathbf{c}^{(p)} = (c_i^{(p)})_{i \geq 1}$ .

On the one hand, we observe that

$$\begin{aligned} \sum_{i=1}^{\infty} i^2 \mathbf{B}(i, 1 + 1/p) &= \int_0^1 \left( \sum_{i=1}^{\infty} i^2 x^{i-1} \right) (1-x)^{1/p} dx \\ &= \int_0^1 (1+x)(1-x)^{-3+1/p} dx \\ &= \frac{p}{(1-p)(1-2p)}, \end{aligned}$$

so that

$$\sum_{i=1}^{\infty} i^2 c_i^{(p)} = \frac{1}{1-2p}. \quad (10)$$

On the other hand, we note the basic identity

$$\mathcal{S}^2(n) = \sum_{j=1}^{\infty} N_j(n)^2 = \sum_{i=1}^{\infty} i^2 C_i(n). \quad (11)$$

Since  $\Gamma(n+2p)/\Gamma(n) \sim n^{2p}$  and  $2p < 1$ , we see from Lemma 2.6 and (11) that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{i=1}^{\infty} i^2 \frac{C_i(n)}{n} \right) = \frac{1}{1-2p}.$$

Thanks to (10), we deduce from the Vitali-Scheffé theorem (see e.g. Theorem 2.8.9 in [9]) that the convergence (9) also holds in  $L^1(\mathbb{P} \otimes \#^2)$ , that is

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{i=1}^{\infty} i^2 \left| \frac{C_i(\ell(n))}{\ell(n)} - c_i^{(p)} \right| \right) = 0.$$

Since the convergence above holds for any (initial) subsequence, our claim is proven.  $\square$

We next point at the following consequence of Lemma 3.1.

**Corollary 3.2** *We have*

$$\lim_{n \rightarrow \infty} \frac{\mathcal{S}^2(n)}{n} = \frac{1}{1-2p} \quad \text{in } L^1(\mathbb{P}),$$

and

$$\lim_{n \rightarrow \infty} \sup_{j \geq 1} \frac{N_j(n)}{\sqrt{n}} = 0 \quad \text{in probability.}$$

**Proof** Observe from (10), (11), and the triangle inequality that

$$\begin{aligned} \mathbb{E} \left( \left| \frac{\mathcal{S}^2(n)}{n} - \frac{1}{1-2p} \right| \right) &= \mathbb{E} \left( \left| \sum_{i=1}^{\infty} i^2 \frac{C_i(n)}{n} - \sum_{i=1}^{\infty} i^2 c_i^{(p)} \right| \right) \\ &\leq \mathbb{E} \left( \sum_{i=1}^{\infty} i^2 \left| \frac{C_i(n)}{n} - c_i^{(p)} \right| \right). \end{aligned}$$

Our first assertion thus follows from Lemma 3.1.

For the second assertion, observe that

$$\sup_{j \geq 1} N_j(n) = \sup\{i \geq 1 : C_i(n) \geq 1\}.$$

We then have for every  $\eta > 0$  arbitrarily small

$$\begin{aligned} \mathbb{P} \left( \sup_{j \geq 1} N_j(n) > \eta \sqrt{n} \right) &= \mathbb{P}(\exists i \geq \eta \sqrt{n} : C_i(n) \geq 1) \\ &\leq \frac{1}{\eta^2 n} \mathbb{E} \left( \sum_{i \geq \eta \sqrt{n}} i^2 C_i(n) \right). \end{aligned}$$

It follows from Lemma 3.1 that the right-hand side converges to 0 as  $n \rightarrow \infty$ , and the proof is now complete.  $\square$

Theorem 1.2(i) now derives immediately from (4), Lemma 2.1(i) and Corollary 3.2 by setting  $\beta_{n,j} = N_j(n)/\sqrt{n}$  for every  $j \geq 1$ .

### 3.2 Critical regime $p = 1/2$

Throughout this section, we assume that the reinforcement parameter is  $p = 1/2$ . Recall from Lemma 2.6 that  $\mathbb{E}(\mathcal{S}^2(n)) \sim n \log n$  as  $n \rightarrow \infty$ . We establish now a stronger version of this estimate.

**Lemma 3.3** *One has*

$$\lim_{n \rightarrow \infty} \frac{\mathcal{S}^2(n)}{n \log n} = 1 \quad \text{almost surely.}$$

**Proof** It has been observed in [7] that, in the study of reinforcement induced by Simon's algorithm, it may be convenient to perform a time-substitution based on a Yule process. We shall use this idea here again, and introduce a standard Yule process  $Y = (Y_t)_{t \geq 0}$ , which we further assume to be independent of the preceding variables. Recall that  $Y$  is a pure birth process in continuous time started from  $Y_0 = 1$  and with birth rate  $n$  from any state  $n \geq 1$ ; in particular, for every function  $f : \mathbb{N} \rightarrow \mathbb{R}$ , say such that  $f(n) = O(n^r)$  for some  $r > 0$ , the process

$$f(Y_t) - \int_0^t (f(Y_s + 1) - f(Y_s)) Y_s ds$$

is a martingale.

Consider the time changed process  $\mathcal{S}^2 \circ Y$ . Applying the observation above to  $f = \mathcal{S}^2$  and then projecting on the natural filtration of  $\mathcal{S}^2 \circ Y$ , the same calculation as in the proof of Lemma 2.6 shows that

$$\begin{aligned} \mathcal{S}^2(Y_t) - \int_0^t \left( \frac{1}{2} + \sum_{i=1}^{\infty} \frac{N_i(Y_s)}{2Y_s} (2N_i(Y_s) + 1) \right) Y_s ds \\ = \mathcal{S}^2(Y_t) - \int_0^t (\mathcal{S}^2(Y_s) + Y_s) ds \end{aligned}$$

is a martingale. By elementary stochastic calculus, the same holds for

$$M_t = e^{-t} \mathcal{S}^2(Y_t) - \int_0^t e^{-s} Y_s ds.$$

We shall now show that  $M$  is bounded in  $L^2(\mathbb{P})$  by checking that its quadratic variation  $[M]_\infty$  has a finite expectation. Plainly,  $M$  is purely discontinuous; its jumps can arise either due to an innovation event (whose instantaneous rate at time  $t$  equals  $\frac{1}{2} Y_{t-}$ ), and then  $\Delta M_t = M_t - M_{t-} = e^{-t}$ , or by a repetition of the  $j$ -th item for some  $j \geq 1$  (whose instantaneous rate at time  $t$  equals  $\frac{1}{2} N_j(Y_{t-})$ ), and then  $\Delta M_t = e^{-t} (2N_j(Y_{t-}) + 1)$ . We thus find by a standard calculation of compensation

that

$$\begin{aligned}\mathbb{E}([M]_\infty) &= \mathbb{E}\left(\sum_{t>0}(\Delta M_t)^2\right) \\ &= \mathbb{E}\left(\int_0^\infty e^{-2t}\left(\frac{1}{2}Y_t + \frac{1}{2}\sum_{j\geq 1}N_j(Y_t)(2N_j(Y_t)+1)^2\right)dt\right) \\ &= \int_0^\infty \mathbb{E}\left(Y_t + 2\sum_{j\geq 1}(N_j(Y_t)^3 + N_j(Y_t)^2)\right)e^{-2t}dt.\end{aligned}$$

First, recall that  $Y_t$  has the geometric distribution with parameter  $e^{-t}$ , in particular  $\int_0^\infty \mathbb{E}(Y_t)e^{-2t}dt = 1$ . Second,  $\sum_{j\geq 1}N_j(Y_t)^2 = \mathcal{S}^2(Y_t)$ , and since  $\mathbb{E}(\mathcal{S}^2(n)) \sim n \log n$  (see Lemma 2.6) and the processes  $S$  and  $Y$  are independent, we have also

$$\int_0^\infty \mathbb{E}\left(\sum_{j\geq 1}N_j(Y_t)^2\right)e^{-2t}dt < \infty.$$

Third, consider  $T(Y_t) = \sum_{j\geq 1}N_j(Y_t)^3$ . By calculations similar to those for  $M_t$ , one sees that the process

$$e^{-3t/2}T(Y_t) - \int_0^t e^{-3s/2}(Y_s + \mathcal{S}^2(Y_s))ds, \quad t \geq 0$$

is a local martingale. Just as above, one readily checks that

$$\int_0^\infty e^{-3s/2}\mathbb{E}(Y_s + \mathcal{S}^2(Y_s))ds < \infty,$$

and hence  $\mathbb{E}(T(Y_t)) = O(e^{3t/2})$ . As a consequence,

$$\int_0^\infty \mathbb{E}\left(\sum_{j\geq 1}N_j(Y_t)^3\right)e^{-2t}dt < \infty,$$

and putting the pieces together, we have checked that  $\mathbb{E}([M]_\infty) < \infty$ .

We now know that  $\lim_{t\rightarrow\infty} M_t = M_\infty$  a.s. and in  $L^2(\mathbb{P})$ , and recall the classical feature that  $\lim_{t\rightarrow\infty} e^{-t}Y_t = W$  a.s., where  $W$  has the standard exponential distribution. In particular  $\int_0^t e^{-s}Y_s ds \sim tW$  as  $t \rightarrow \infty$ , so that

$$\mathcal{S}^2(Y_t) = te^t W + o(e^t), \quad \text{a.s.}$$

Using again  $Y_t = e^t W + o(e^t)$ , we conclude that  $\mathcal{S}^2(n) = n \log n + O(n)$  a.s., which implies our claim.  $\square$

**Remark 3.4** The first part of Corollary 3.2 and Lemma 3.3 seem to be of the same nature. Actually, one can also establish the former by adapting the proof of the latter, therefore circumventing the appeal to Lemma 2.5. There is nonetheless a fundamental difference between these two results: although the microscopic clusters (i.e. of size  $O(1)$ ) determine the asymptotic behavior of  $\mathcal{S}^2(n)$  in the sub-critical case, they have no impact in the critical case as it is seen from Lemma 2.5.

Thanks to Lemmas 2.1(ii) and 3.3, the following statement is the final piece of the proof of Theorem 1.2(ii).

**Lemma 3.5** *One has*

$$\lim_{n \rightarrow \infty} \sup_{j \geq 1} \frac{N_j(n)}{\sqrt{n \log n}} = 0 \quad \text{in probability.}$$

**Proof** We shall show that there is some numerical constant  $b$  such that

$$\mathbb{E}(N_j(n)^3) \leq b(n/j)^{3/2} \quad \text{for all } j, n \geq 1. \quad (12)$$

Then, by Markov's inequality, we have that for any  $\eta > 0$

$$\mathbb{P}\left(N_j(n) > \sqrt{\eta n \log n}\right) \leq b(\eta n \log n)^{-3/2} (n/j)^{3/2},$$

and by the union bound

$$\mathbb{P}\left(\exists j \leq n : N_j(n) > \sqrt{\eta n \log n}\right) \leq b(\eta \log n)^{-3/2} \sum_{j \geq 1} j^{-3/2},$$

which proves our claim.

For  $i = 1, 2, 3$ , set  $a_i(n) = \Gamma(n + i/2) / \Gamma(n)$ , so  $a_i(n) \sim n^{i/2}$  and actually  $a_2(n) = n$ . Recall that  $i(n)$  denotes the number of innovations up to the  $n$ -step of Simon's algorithm. Take any  $j \geq 1$  and, just as in the proof of Lemma 2.6, observe that on the event  $i(n) \geq j$ , one has

$$\begin{aligned} \mathbb{E}\left(\frac{N_j(n+1)}{a_1(n+1)} \mid \mathcal{F}_n\right) &= \frac{N_j(n)}{a_1(n)}, \\ \mathbb{E}\left(\frac{N_j(n+1)^2}{a_2(n+1)} \mid \mathcal{F}_n\right) &= \frac{N_j(n)^2}{a_2(n)} + \frac{N_j(n)}{2na_2(n+1)}, \\ \mathbb{E}\left(\frac{N_j(n+1)^3}{a_3(n+1)} \mid \mathcal{F}_n\right) &= \frac{N_j(n)^3}{a_3(n)} + \frac{3N_j(n)^2}{2na_3(n+1)} + \frac{N_j(n)}{2na_3(n+1)}. \end{aligned}$$

The trivial bound  $i(j) \leq j$  then yields for any  $n \geq j$

$$\mathbb{E}(N_j(n)) \leq a_1(n)/a_1(j) \leq b_1 \sqrt{n/j}.$$



Then we have

$$\mathbb{E}(N_j(n)^2) \leq \frac{a_2(n)}{a_2(j)} + \frac{b_1}{2na_2(n)} \sum_{k=j}^n \sqrt{k/j} \leq b_2 n/j,$$

and finally also

$$\mathbb{E}(N_j(n)^3) \leq \frac{a_3(n)}{a_3(j)} + \frac{3b_2}{2na_3(n)} \sum_{k=j}^n k/j + \frac{b_1}{2na_3(n)} \sum_{k=j}^n \sqrt{k/j} \leq b_3(n/j)^{3/2},$$

where  $b_1$ ,  $b_2$  and  $b_3$  are numerical constants. This establishes (12) and completes the proof.  $\square$

### 3.3 Supercritical regime $p > 1/2$

Throughout this section, we assume that the reinforcement parameter satisfies  $p > 1/2$ . We first point at the following strengthening of Lemma 2.2 (in particular, recall the notation (5) there).

**Corollary 3.6** *We have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{j=1}^{\infty} \left| \frac{N_j(n)}{n^p} - X_j^{(p)} \right|^2 \right) = 0.$$

This result has been already observed by Businger, see Equation (6) in [10]. For the sake of completeness, we present here an alternative and shorter proof along the same line as for Lemma 3.1.

**Proof** We view  $\mathbf{X}^{(p)} = (X_j^{(p)})_{j \geq 1}$  and  $\mathbf{N}(n) = (N_j(n))_{j \geq 1}$  for each  $n \geq 1$  as functions on the space  $\Omega \times \mathbb{N}$  endowed with the product measure  $\mathbb{P} \otimes \#$ , where  $\#$  denotes the counting measure on  $\mathbb{N}$ . Since we already know from Lemma 2.2 that  $n^{-p}\mathbf{N}(n)$  converges as  $n \rightarrow \infty$  to  $\mathbf{X}^{(p)}$  almost everywhere, in order to establish our claim, it suffices to verify that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{j=1}^{\infty} \frac{N_j(n)^2}{n^{2p}} \right) = \mathbb{E} \left( \sum_{j=1}^{\infty} (X_j^{(p)})^2 \right);$$

see e.g. Proposition 4.7.30 in [9].

Recall from Lemma 2.6 that

$$\mathbb{E} \left( \sum_{j=1}^{\infty} \frac{N_j(n)^2}{n^{2p}} \right) = \frac{\Gamma(n+2p)}{n^{2p}\Gamma(n)} \sum_{i=1}^n \frac{\Gamma(i)}{\Gamma(i+2p)}.$$

On the one hand, we know that

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+2p)}{n^{2p}\Gamma(n)} = 1,$$

and on the other hand, we recall from (7) that

$$\sum_{i=1}^{\infty} \frac{\Gamma(i)}{\Gamma(i+2p)} = \frac{1}{(2p-1)\Gamma(2p)}.$$

We conclude from Lemma 2.2 that indeed

$$\lim_{n \rightarrow \infty} \mathbb{E} \left( \sum_{j=1}^{\infty} \frac{N_j(n)^2}{n^{2p}} \right) = \frac{1}{(2p-1)\Gamma(2p)} = \mathbb{E} \left( \sum_{j=1}^{\infty} (X_j^{(p)})^2 \right)$$

and the proof is complete.  $\square$

Theorem 1.2 can now be deduced from (4), Lemma 2.1(ii), and Corollary 3.6.

## 4 Relation to step reinforced random walks

It is interesting to combine Donsker's Theorem with the continuous mapping theorem; notably considering the overall supremum of paths yields the well-known Kolmogorov–Smirnov test. In this direction, linear mappings of the type  $\omega \mapsto \int_{[0,1]} \omega(x)m(dx)$ , where  $m$  is some finite measure on  $[0, 1]$ , are amongst the simplest functionals on  $\mathbb{D}$ . Writing  $\bar{m}(x) = m((x, 1])$  for the tail distribution function and  $\mu = \int_{[0,1]} xm(dx)$  for the mean, this leads us to consider the variables

$$\xi_j = \bar{m}(U_j) - \mu \quad \text{and} \quad \hat{\xi}_j = \bar{m}(\hat{U}_j) - \mu \quad \text{for } j \geq 1.$$

So  $(\xi_j)_{j \geq 1}$  is an i.i.d. sequence and  $(\hat{\xi}_j)_{j \geq 1}$  can be viewed as the reinforced sequence resulting from Simon's algorithm. All these variables have the same distribution, they are bounded and centered with variance

$$\varsigma^2 = \int_{[0,1]} \int_{[0,1]} (x \wedge y - xy)m(dx)m(dy) = \text{Var} \left( \int_{[0,1]} G(x)m(dx) \right),$$

where  $G = (G(x))_{0 \leq x \leq 1}$  is a Brownian bridge. In this setting, we have

$$\int_{[0,1]} \hat{G}_n(x)m(dx) = \frac{\hat{\xi}_1 + \cdots + \hat{\xi}_n}{\sqrt{n}}.$$

The process of the partial sums

$$\hat{S}(n) = \hat{\xi}_1 + \cdots + \hat{\xi}_n, \quad n \geq 0$$

is called a step reinforced random walk. We now immediately deduce from Theorem 1.2 and the continuous mapping theorem that its asymptotic behavior is given by:

(i) if  $p < 1/2$ , then

$$n^{-1/2} \hat{S}(n) \implies \mathcal{N}(0, \varsigma^2/(1-2p));$$

(ii) if  $p = 1/2$ , then

$$(n \log n)^{-1/2} \hat{S}(n) \implies \mathcal{N}(0, \varsigma^2);$$

(iii) if  $p > 1/2$ , then

$$\lim_{n \rightarrow \infty} n^{-p} \hat{S}(n) = \sum_{j \geq 1} \xi_j X_j^{(p)} \quad \text{in probability,}$$

where  $\mathbf{X}^{(p)} = (X_j^{(p)})_{j \geq 1}$  has been defined in Lemma 2.2 and is independent of the  $\xi_j$ .

Although this argument only enables us to deal with real bounded random variables  $\xi$ , we stress that more generally, the assertions (i), (ii) and (iii) still hold when the generic step  $\xi$  is an arbitrary square integrable and centered variable in  $\mathbb{R}^d$  (for  $d \geq 2$ ,  $\varsigma^2$  is then of course the covariance matrix of  $\xi$ ). Specifically, (i) follows from the invariance principle for step reinforced random walks (see Theorem 3.3 in [7]), whereas (iii) is Theorem 3.2 in the same work; see also [6]. In the critical case  $p = 1/2$ , (ii) can be deduced from the basic identity

$$\hat{S}(n) = \sum_{j=1}^{\infty} N_j(n) \xi_j,$$

the Lévy–Lindeberg theorem (see e.g. Theorem 5.2 of Chapter VII in [17]), and Lemmas 3.3 and 3.5.

In this vein, we mention that when  $\xi$  has the Bernoulli distribution, (i–iii) are due originally to Heyde [16] in the setting of the so-called correlated Bernoulli processes, see also [18, 29, 30]. These results have also appeared more recently in the framework of the so-called elephant random walk, a random walk with memory which has been introduced by Schütz and Trimper [27]. See notably [2, 3, 11, 12], and also [4, 5, 14, 15] and references therein for some further developments. We mention that Kürsten [23] first pointed at the role of Bernoulli bond percolation on random recursive trees in this framework, see also [10]. It is moreover interesting to recall that, for the elephant random walk, Kubota and Takei [22] have established that the fluctuations corresponding to (iii) are Gaussian. Whether or not the same holds for general step reinforced random walks is still open; this also suggests that for  $p > 1/2$ , Theorem

**1.2** (iii) might be refined and yield a second order weak limit theorem involving again a Brownian bridge in the limit.

**Funding** Open access funding provided by University of Zurich.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Baur, E., Bertoin, J.: The fragmentation process of an infinite recursive tree and Ornstein–Uhlenbeck type processes. *Electron. J. Probab.* **20**, 20 (2015)
2. Baur, E., Bertoin, J.: Elephant random walks and their connection to Pólya-type urns. *Phys. Rev. E* **94**, 052134 (2016)
3. Bercu, B.: A martingale approach for the elephant random walk. *J. Phys. A* **51**(1), 015201 (2018)
4. Bercu, B., Laulin, L.: On the multi-dimensional elephant random walk. *J. Stat. Phys.* **175**(6), 1146–1163 (2019)
5. Bertengui, M.: Functional limit theorems for the multi-dimensional elephant random walk. [arXiv:2004.02004](https://arxiv.org/abs/2004.02004)
6. Bertoin, J.: Scaling exponents of step-reinforced random walks. [hal-02480479](https://arxiv.org/abs/2004.04079)
7. Bertoin, J.: Universality of noise reinforced Brownian motions. In: In and Out of Equilibrium 3, celebrating Vladas Sidoravicius (to appear), *Progress in Probability*. Birkhäuser. [arXiv:2002.09166](https://arxiv.org/abs/2002.09166)
8. Billingsley, P.: *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics, 2nd edn. Wiley, New York (1999)
9. Bogachev, V.I.: *Measure Theory*, vol. I, II. Springer, Berlin (2007)
10. Businger, S.: The shark random swim (Lévy flight with memory). *J. Stat. Phys.* **172**(3), 701–717 (2018)
11. Coletti, C.F., Gava, R., Schütz, G.M.: Central limit theorem and related results for the elephant random walk. *J. Math. Phys.* **58**(5), 053303 (2017). 8
12. Coletti, C.F., Gava, R., Schütz, G.M.: A strong invariance principle for the elephant random walk. *J. Stat. Mech. Theory Exp.* **12**, 123207 (2017). 8
13. Drmota, M.: *Random Trees: An Interplay Between Combinatorics and Probability*, 1st edn. Springer, Berlin (2009)
14. González-Navarrete, M.: Multidimensional walks with random tendency. [arXiv:2004.04033](https://arxiv.org/abs/2004.04033)
15. Guevara, V.H.V.: On the almost sure central limit theorem for the elephant random walk. *J. Phys. A: Math. Theor.* **52**(47), 475201 (2019)
16. Heyde, C.: Asymptotics and criticality for a correlated Bernoulli process. *Aust. N. Z. J. Stat.* **46**(1), 53–57 (2004)
17. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*. Grundlehren der mathematischen Wissenschaften. Springer, Berlin (2002)
18. James, B., James, K., Qi, Y.: Limit theorems for correlated Bernoulli random variables. *Stat. Probab. Lett.* **78**(15), 2339–2345 (2008)
19. Kallenberg, O.: Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **27**, 23–36 (1973)
20. Kallenberg, O.: Path properties of processes with independent and interchangeable increments. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **28**, 257–271 (1973/74)
21. Kesten, H.: The critical probability of bond percolation on the square lattice equals  $\frac{1}{2}$ . *Commun. Math. Phys.* **74**(1), 41–59 (1980)
22. Kubota, N., Takei, M.: Gaussian fluctuation for superdiffusive elephant random walks. *J. Stat. Phys.* **177**(6), 1157–1171 (2019)

23. Kürsten, R.: Random recursive trees and the elephant random walk. *Phys. Rev. E* **93**(3), 032111 (2016). 11
24. Möhle, M.: The Mittag–Leffler process and a scaling limit for the block counting process of the Bolthausen–Sznitman coalescent. *ALEA Lat. Am. J. Probab. Math. Stat.* **12**(1), 35–53 (2015)
25. Pachon, A., Polito, F., Sacerdote, L.: Random graphs associated to some discrete and continuous time preferential attachment models. *J. Stat. Phys.* **162**(6), 1608–1638 (2016)
26. Rio, E.: *Asymptotic Theory of Weakly Dependent Random Processes*. Probability Theory and Stochastic Modelling, vol. 80. Springer, Berlin (2017). Translated from the 2000 French edition
27. Schütz, G.M., Trimper, S.: Elephants can always remember: exact long-range memory effects in a non-Markovian random walk. *Phys. Rev. E* **70**, 045101 (2004)
28. Simon, H.A.: On a class of skew distribution functions. *Biometrika* **42**(3/4), 425–440 (1955)
29. Wu, L., Qi, Y., Yang, J.: Asymptotics for dependent Bernoulli random variables. *Stat. Probab. Lett.* **82**(3), 455–463 (2012)
30. Zhang, Y., Zhang, L.-X.: On the almost sure invariance principle for dependent Bernoulli random variables. *Stat. Probab. Lett.* **107**, 264–271 (2015)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.